

RECEIVED

JUN 12 2002

TECH CENTER 1600/2900

1653

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:08

Input Set : A:\140p1-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

SEQUENCE LISTING

ENTERED

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: He, Wei-Wu et al.

8 (ii) TITLE OF INVENTION: Interleukin-1 Beta Converting Enzyme.

9 Like Apoptosis Protease 3 and 4

11 (iii) NUMBER OF SEQUENCES: 12

13 (iv) CORRESPONDENCE ADDRESS:

14 (A) ADDRESSEE: Human Genome Sciences, Inc.

15 (B) STREET: 9410 Key West Ave.

16 (C) CITY: Rockville

17 (D) STATE: MD

18 (E) COUNTRY: USA

19 (F) ZIP: 20850

21 (v) COMPUTER READABLE FORM:

22 (A) MEDIUM TYPE: Floppy disk

23 (B) COMPUTER: IBM PC compatible

24 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

27 (vi) CURRENT APPLICATION DATA:

C--> 28 (A) APPLICATION NUMBER: US/09/613,508

C--> 29 (B) FILING DATE: 10-Jul-2000

30 (C) CLASSIFICATION:

32 (vii) PRIOR APPLICATION DATA:

33 (A) APPLICATION NUMBER: US 08/334,251

34 (B) FILING DATE: 11-NOV-1994

36 (viii) ATTORNEY/AGENT INFORMATION:

37 (A) NAME: Brookes, A. Anders

38 (B) REGISTRATION NUMBER: 36,373

39 (C) REFERENCE/DOCKET NUMBER: PF140P1

41 (ix) TELECOMMUNICATION INFORMATION:

42 (A) TELEPHONE: 301-309-8504

43 (B) TELEFAX: 301-309-8439

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 1371 base pairs

50 (B) TYPE: nucleic acid

51 (C) STRANDEDNESS: single

52 (D) TOPOLOGY: linear

54 (ii) MOLECULE TYPE: DNA (genomic)

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61 GCACGAGAAA CTTTGCTGTG CGCGTTCTCC CGCGCGCGGG CTCAACTTTG TAGAGCGAGG	60
63 GGCCAACTTG GCAGAGCGCG CGGCCAGCTT TGCAGAGAGC GCCCTCCAGG GACTATGCGT	120
65 GCGGGGACAC GGGTCGCTTT GGGCTCTTCC ACCCTGCGG AGCGCACTAC CCCGAGCCAG	180

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:08

Input Set : A:\140pl-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

```

67 GGGCGGTGCA AGCCCCGCCC GGCCCTACCC AGGGCGGGCTC CTCCCTCCGC AGCGCCGAGA      240
69 CTTTTAGTTT CGCTTTTCGCT AAAGGGGGCCC CAGACCCTTG CTGCGGAGCG ACGGAGAGAG      300
71 ACTGTGCCAG TCCCAGCCGC CCTACCGCCG TGGGAACGAT GGCAGATGAT TCAGGGCTGT      360
73 ATTGAAGAGC AGGGGGTTGA GGATTCAGCA AATGAAGATT CAGTGGATGC TAAGCCAGAC      420
75 CGGTCCTCGT TTGTACCGTC CCTCTTCAGT AAGAAGAAGA AAAATGTCAC CATGCGATCC      480
77 ATCAAGACCA CCCGGGACCG AGTGCCTACA TATCAGTACA ACATGAATTT TGAAAAGCTG      540
79 GGCAAATGCA TCATAATAAA CAACAAGAAC TTTGATAAAG TGACAGGTAT GGGCGTTTCGA      600
81 AACGGAACAG ACAAAGATGC CGAGGCGCTC TTCAAGTGCT TCCGAAGCCT GGGTTTTGAC      660
83 GTGATTGTCT ATAATGACTG CTCTGTGTCC AAGATGCAAG ATCTGCTTAA AAAAGCTTCT      720
85 GAAGAGGACC ATACAAATGC CGCCTGCTTC GCCTGCATCC TCTTAAGCCA TGGAGAAGAA      780
87 AATGTAATTT ATGGGAAAGA TGGTGTCA CAATAAAGG ATTTGACAGC CCACTTTAGG      840
89 GGGGATAGAT GCAAAACCTT TTAGAGAAA CCCAACTCT TCTTCATTCA GGCTTGCCGA      900
91 GGGACCGAGC TTGATGATGG CATCCAGGCC GACTCGGGGC CCATCAATGA CACAGATGCT      960
93 AATCCTCGAT ACAAGATCCC AGTGAAGCT GACTTCCTCT TCGCCTATTC CACGGTTCCA      1020
95 GGCTATTACT CGTGGAGGAG CCCAGGAAGA GGCTCCTGGT TTGTGCAAGC CCTCTGCTCC      1080
97 ATCCTGGAGG AGCACGAAA AGACCTGGA ATCATGCAAA TCCTCCACCA GGGTGAATGA      1140
99 CAGAGTTGCC AGGCACTTTG AGTCTCAGTC TGATGACCCA CACTTCCATG AGAAGAAGCA      1200
101 GATCCCCTGT GTGGTCTCCA TGCTACCAA GGAACCTCTAC TTCAGTCAAT AGCCATATCA      1260
103 GGGGTACATT CTAGCTGAGA AGCAATGGGT CACTCATTA TGAATCACAT TTTTTTATGC      1320
105 TCTTGAAATA TTCAGAAATT CTCCAGGATT TTAATTTTCA GAAAATGTAT T      1371

```

107 (2) INFORMATION FOR SEQ ID NO: 2:

109 (i) SEQUENCE CHARACTERISTICS:

110 (A) LENGTH: 303 amino acids

111 (B) TYPE: amino acid

112 (C) STRANDEDNESS: single

113 (D) TOPOLOGY: linear

115 (ii) MOLECULE TYPE: protein

120 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

122 Met Ala Asp Asp Gln Gly Cys Ile Glu Glu Gln Gly Val Glu Asp Ser
123 1 5 10 15
125 Ala Asn Glu Asp Ser Val Asp Ala Lys Pro Asp Arg Ser Ser Phe Val
126 20 25 30
128 Pro Ser Leu Phe Ser Lys Lys Lys Asn Val Thr Met Arg Ser Ile
129 35 40 45
131 Lys Thr Thr Arg Asp Arg Val Pro Thr Tyr Gln Tyr Asn Met Asn Phe
132 50 55 60
134 Glu Lys Leu Gly Lys Cys Ile Ile Ile Asn Asn Lys Asn Phe Asp Lys
135 65 70 75 80
137 Val Thr Gly Met Gly Val Arg Asn Gly Thr Asp Lys Asp Ala Glu Ala
138 85 90 95
140 Leu Phe Lys Cys Phe Arg Ser Leu Gly Phe Asp Val Ile Val Tyr Asn
141 100 105 110
143 Asp Cys Ser Cys Ala Lys Met Gln Asp Leu Leu Lys Lys Ala Ser Glu
144 115 120 125
146 Glu Asp His Thr Asn Ala Ala Cys Phe Ala Cys Ile Leu Leu Ser His
147 130 135 140
149 Gly Glu Glu Asn Val Ile Tyr Gly Lys Asp Gly Val Thr Pro Ile Lys
150 145 150 155 160
152 Asp Leu Thr Ala His Phe Arg Gly Asp Arg Cys Lys Thr Leu Leu Glu

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:08

Input Set : A:\140pl-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

```

153          165          170          175
155      Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Arg Gly Thr Glu Leu Asp
156          180          185          190
158      Asp Ala Ile Gln Ala Asp Ser Gly Pro Ile Asn Asp Thr Asp Ala Asn
159          195          200          205
161      Pro Arg Tyr Lys Ile Pro Val Glu Ala Asp Phe Leu Phe Ala Tyr Ser
162          210          215          220
164      Thr Val Pro Gly Tyr Tyr Ser Trp Arg Ser Pro Gly Arg Gly Ser Trp
165          225          230          235          240
167      Phe Val Gln Ala Leu Cys Ser Ile Leu Glu Glu His Gly Lys Glu Leu
168          245          250          255
170      Glu Ile Met Gln Ile Leu Thr Arg Val Asn Asp Arg Val Ala Arg His
171          260          265          270
173      Phe Glu Ser Gln Ser Asp Asp Pro His Phe His Glu Lys Lys Gln Ile
174          275          280          285
176      Pro Cys Val Val Ser Met Leu Thr Lys Glu Leu Tyr Phe Ser Gln
177          290          295          300
179 (2) INFORMATION FOR SEQ ID NO: 3:
181      (i) SEQUENCE CHARACTERISTICS:
182          (A) LENGTH: 1159 base pairs
183          (B) TYPE: nucleic acid
184          (C) STRANDEDNESS: single
185          (D) TOPOLOGY: linear
187      (ii) MOLECULE TYPE: DNA (genomic)
192      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
194      GCACGAGCGG ATGGGTGCTA TTGTGAGGCG GTTGTAGAAG AGTTTCGTGA GTGCTCGCAG      60
196      CTCATACCTG TGGCTGTGTA TCCGTGGCCA CAGCTGGTTG GCGTCGCCTT GAAATCCCGAG      120
198      GCCGTGAGGA GTTAGCGAGC CCTGCTCACA CTCGGCGCTC TGGTTTTTCGG TGGGTGTGCC      180
200      CTGCACCTGC CTCTTCCCGC ATTCTCATTA ATAAAGGTAT CCATGGAGAA CACTGAAAAC      240
202      TCAGTGGATT CAAAATCCAT TAAAAATTTG GAACCAAAGA TCATACATGG AAGCGAATCA      300
204      ATGGACTCTG GAATATCCCT GGACAACAGT TATAAAATGG ATTATCCTGA GATGGGTTTA      360
206      TGTATAATAA TTAATAATAA GAATTTTCAT AAAAGCACTG GAATGACATC TCGGTCTGGT      420
208      ACAGATGTCG ATGCAGCAAA CCTCAGGGAA ACATTCAGAA ACTTGAAATA TGAAGTCAGG      480
210      AATAAAAAATG ATCTTACACG TGAAGAAATT GTGGAATTGA TGGGTGATGT TTCTAAAGAA      540
212      GATCACAGCA AAAGGAGCAG TTTTGTTTGT GTGCTTCTGA GCCATGGTGA AGAAGGAATA      600
214      ATTTTTGGAA CAAATGGACC TGTTGACCTG AAAAAAATAA CAAACTTTTT CAGAGGGGAT      660
216      CGTTGTAGAA GTCTAACTGG AAAACCCAAA CTTTTCATTA TTCAGGCCTG CCGTGGTACA      720
218      GAACTGGACT GTGGCATTGA GACAGACAGT GGTGTTGATG ATGACATGGC GTGTCATAAA      780
220      ATACCAGTGG AGGCCGACTT CTTGTATGCA TACTCCACAG CACCTGGTTA TTATTCCTGG      840
222      CGAAATTCAA AGGATGGCTC CTGGTTCATC CAGTCGCTTT GTGCCATGCT GAAACAGTAT      900
224      GCCGACAAGC TTGAATTTAT GCACATTCTT ACCCGGGTTA ACCGAAAGGT GGCAACAGAA      960
226      TTTGAGTCCT TTTCTTTTGA CGTACTTTT CATGCAAAGA AACAGATTCC ATGTATTGTT      1020
228      TCCATGCTCA CAAAAGAACT CTATTTTAT CACTAAAGAA ATGGTTGGTT GGTGGTTTTT      1080
230      TTTAGTTTGT ATGCCAAGTG AGAAGATGGT ATATTTGGGT ACTGTATTTC CCTCTCATTG      1140
232      GGGACCTACT CTCATGCTG      1159
234 (2) INFORMATION FOR SEQ ID NO: 4:
236      (i) SEQUENCE CHARACTERISTICS:
237          (A) LENGTH: 277 amino acids
238          (B) TYPE: amino acid

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:08

Input Set : A:\140pl-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

```

239      (C) STRANDEDNESS: single
240      (D) TOPOLOGY: linear
242      (ii) MOLECULE TYPE: protein
247      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
249      Met Glu Asn Thr Glu Asn Ser Val Asp Ser Lys Ser Ile Lys Asn Leu
250      1          5          10          15
252      Glu Pro Lys Ile Ile His Gly Ser Glu Ser Met Asp Ser Gly Ile Ser
253      20          25          30
255      Leu Asp Asn Ser Tyr Lys Met Asp Tyr Pro Glu Met Gly Leu Cys Ile
256      35          40          45
258      Ile Ile Asn Asn Lys Asn Phe His Lys Ser Thr Gly Met Thr Ser Arg
259      50          55          60
261      Ser Gly Thr Asp Val Asp Ala Ala Asn Leu Arg Glu Thr Phe Arg Asn
262      65          70          75          80
264      Leu Lys Tyr Glu Val Arg Asn Lys Asn Asp Leu Thr Arg Glu Glu Ile
265      85          90          95
267      Val Glu Leu Met Arg Asp Val Ser Lys Glu Asp His Ser Lys Arg Ser
268      100         105         110
270      Ser Phe Val Cys Val Leu Leu Ser His Gly Glu Glu Gly Ile Ile Phe
271      115         120         125
273      Gly Thr Asn Gly Pro Val Asp Leu Lys Lys Ile Thr Asn Phe Phe Arg
274      130         135         140
276      Gly Asp Arg Cys Arg Ser Leu Thr Gly Lys Pro Lys Leu Phe Ile Ile
277      145         150         155         160
279      Gln Ala Cys Arg Gly Thr Glu Leu Asp Cys Gly Ile Glu Thr Asp Ser
280      165         170         175
282      Gly Val Asp Asp Asp Met Ala Cys His Lys Ile Pro Val Glu Ala Asp
283      180         185         190
285      Phe Leu Tyr Ala Tyr Ser Thr Ala Pro Gly Tyr Tyr Ser Trp Arg Asn
286      195         200         205
288      Ser Lys Asp Gly Ser Trp Phe Ile Gln Ser Leu Cys Ala Met Leu Lys
289      210         215         220
291      Gln Tyr Ala Asp Lys Leu Glu Phe Met His Ile Leu Thr Arg Val Asn
292      225         230         235         240
294      Arg Lys Val Ala Thr Glu Phe Glu Ser Phe Ser Phe Asp Ala Thr Phe
295      245         250         255
297      His Ala Lys Lys Gln Ile Pro Cys Ile Val Ser Met Leu Thr Lys Glu
298      260         265         270
300      Leu Tyr Phe Tyr His
301      275
303      (2) INFORMATION FOR SEQ ID NO: 5:
305      (i) SEQUENCE CHARACTERISTICS:
306          (A) LENGTH: 31 base pairs
307          (B) TYPE: nucleic acid
308          (C) STRANDEDNESS: single
309          (D) TOPOLOGY: linear
311      (ii) MOLECULE TYPE: DNA (genomic)
316      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
318      GATCGGATCC ATGCGTGCGG GGACACGGGT C

```

31

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:08

Input Set : A:\140p1-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

```

320 (2) INFORMATION FOR SEQ ID NO: 6:
322     (i) SEQUENCE CHARACTERISTICS:
323         (A) LENGTH: 31 base pairs
324         (B) TYPE: nucleic acid
325         (C) STRANDEDNESS: single
326         (D) TOPOLOGY: linear
328     (ii) MOLECULE TYPE: DNA (genomic)
333     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
335 GTACTCTAGA TCATTCACCC TGGTGGAGGA T 31
337 (2) INFORMATION FOR SEQ ID NO: 7:
339     (i) SEQUENCE CHARACTERISTICS:
340         (A) LENGTH: 31 base pairs
341         (B) TYPE: nucleic acid
342         (C) STRANDEDNESS: single
343         (D) TOPOLOGY: linear
345     (ii) MOLECULE TYPE: DNA (genomic)
350     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
352 GATCGGATCC ATGGAGAACA CTGAAACTC A 31
354 (2) INFORMATION FOR SEQ ID NO: 8:
356     (i) SEQUENCE CHARACTERISTICS:
357         (A) LENGTH: 31 base pairs
358         (B) TYPE: nucleic acid
359         (C) STRANDEDNESS: single
360         (D) TOPOLOGY: linear
362     (ii) MOLECULE TYPE: DNA (genomic)
367     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
369 GTACTCTAGA TTAGTGATAA AAATAGAGTT C 31
371 (2) INFORMATION FOR SEQ ID NO: 9:
373     (i) SEQUENCE CHARACTERISTICS:
374         (A) LENGTH: 22 base pairs
375         (B) TYPE: nucleic acid
376         (C) STRANDEDNESS: single
377         (D) TOPOLOGY: linear
379     (ii) MOLECULE TYPE: DNA (genomic)
384     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
386 GACTATGCGT GCGGGGACAC GG 22
388 (2) INFORMATION FOR SEQ ID NO: 10:
390     (i) SEQUENCE CHARACTERISTICS:
391         (A) LENGTH: 53 base pairs
392         (B) TYPE: nucleic acid
393         (C) STRANDEDNESS: single
394         (D) TOPOLOGY: linear
396     (ii) MOLECULE TYPE: DNA (genomic)
401     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
403 AATCAAGCGT AGTCTGGGAC GTCGTATGGG TATTCACCCT GGTGGAGGAT TTG 53
405 (2) INFORMATION FOR SEQ ID NO: 11:
407     (i) SEQUENCE CHARACTERISTICS:
408         (A) LENGTH: 21 base pairs
409         (B) TYPE: nucleic acid

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/613,508

DATE: 05/31/2002

TIME: 08:28:09

Input Set : A:\140p1-sl.txt

Output Set: N:\CRF3\05312002\I613508.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]